



TITLE:

# 長期記憶モデルにおける二・三の問題(統計的決定方式の最適性の研究)

AUTHOR(S):

矢島, 美寛

---

CITATION:

矢島, 美寛. 長期記憶モデルにおける二・三の問題(統計的決定方式の最適性の研究). 数理解析研究所講究録 1990, 723: 1-14

ISSUE DATE:

1990-05

URL:

<http://hdl.handle.net/2433/101861>

RIGHT:

## 長期記憶モデルにおける二・三の問題

和歌山大・経 矢島美寛 (Yoshihiro Yajima)

1. 序

観測値間の系列相関の強へ、つわゆる「長期記憶モデル」においては、従来の ARMA モデル等、系列相関の弱へモデルについて成立した統計的性質をそのまま適用することはできない。ここではまず一例として、誤差項が長期記憶モデルにしたがう回帰モデルの最小二乗推定量 (LSE) について論じる。後に今後解決すべきと思われる問題のうち、いくつかを紹介する。

2. 回帰モデルモデル

$$y_t = X_t' \beta + \varepsilon_t$$

$$X_t = (x_{t1}, \dots, x_{tk})' \quad \text{説明変数}$$

$$\beta = (\beta_1, \dots, \beta_k)' \quad \text{未知パラメータ}$$

$$\{\varepsilon_t\}: \text{誤差項 (定常過程)} \quad E\varepsilon_t = 0$$

$$f(\lambda) = f^*(\lambda) / |1 - e^{i\lambda}|^{2\alpha}, \quad 0 < \alpha < 1/2$$

$f^*(\lambda)$  : positive continuous

以下では  $f(\lambda)$  を  $\{\varepsilon_t\}$  のスペクトル密度とする.  $f(\lambda)$  は  $\lambda \rightarrow 0$  のとき無限大に発散する. したがって低周波数の波が優勢な密度である. この性質に因んで、 $\{\varepsilon_t\}$  は「長期記憶モデル」にしようという.

### Grenander's conditions

$$a_{ij}^T(h) = \frac{1}{n} \sum_{t=1}^{T-h} x_{t+h,i} x_{t,j}, \quad h=0, 1, \dots,$$

$$\bar{a} = \frac{1}{2} \sum_{t=1-h}^T x_{t+h,i} x_{t,j}, \quad h=0, -1, \dots.$$

(Gr.1)  $a_{ii}^T(0) \rightarrow \infty \quad (T \rightarrow \infty), \quad i=1, \dots, K.$

(Gr.2)  $\lim_{T \rightarrow \infty} x_{T+1,i}^2 / a_{ii}^T(0) = 0, \quad i=1, \dots, K.$

(Gr.3)  $T \rightarrow \infty$  のとき

$$a_{ij}^T(h) / \{a_{ii}^T(0) a_{jj}^T(0)\}^{1/2} (\bar{a} r_{ij}^T(h))$$

の極限が存在.  $i, j=1, 2, \dots, K, \quad h=0, \pm 1, \dots.$

(Gr.4)  $R(0)$  : nonsingular

ここで

$$R(h) = \frac{1}{n} \{p_{ij}(h)\}_{K \times K}$$

$$p_{ij}(h) = \bar{a} \lim_{T \rightarrow \infty} r_{ij}^T(h)$$

このとき  $R(h)$  は

$$R(h) = \int_{-\pi}^{\pi} e^{i\lambda h} dM(\lambda)$$

と表現できる.  $M(\lambda)$  は半正値な増分を持つエルミート行列

## 推定量

$$\hat{\beta}_{T, LSE} = (\hat{X}_T' \hat{X}_T)^{-1} \hat{X}_T' Y_T \quad LSE$$

$$\hat{\beta}_{T, BLUE} = (\hat{X}_T' \Sigma_T^{-1} \hat{X}_T)^{-1} \hat{X}_T' \Sigma_T^{-1} Y_T \quad BLUE$$

ここで

$$Y_T = (y_1, y_2, \dots, y_T)' \quad T \times 1$$

$$\hat{X}_T = (x_{ti}) \quad T \times K$$

$$\Sigma_T = (\sigma_{ij}) \quad T \times T \quad \sigma_{ij} = \sigma_{i-j} (= E \varepsilon_i \varepsilon_j)$$

## LSE の漸近有効性

$\{\varepsilon_t\}$  のスペクトル密度が *positive, continuous* な場合については、BLUE に対して LSE が漸近有効になるための必要十分条件を Grenander (1954) が導出している。説明変数が多項式  $x_{ti} = t^{i-1}$ 、三角関数  $x_{ti} = \sin \omega_i t$  or  $\cos \omega_i t$  ( $\omega_i \neq 0$ ) (両者とも (G.1) ~ (G.4) を満足する) などの代表的な例については LSE は漸近有効となる。しかしながら  $\{\varepsilon_t\}$  が長期記憶モデルにしたがうとき、 $x_{ti} = t^{i-1}$  の場合については、LSE がもはや漸近有効でないことが示されている。(Beran & Künsch (1985), Samoruk & Tagg (1988), Yajima (1988)). ここでは他の一般の説明変数について、LSE の性質を議論する。比較を明瞭にする為、まず Grenander の結果を示す。

## Theorem (Grenander)

$$D_T = \text{diag}(\|X_1\|_T, \|X_2\|_T, \dots, \|X_K\|_T)$$

$$\|X_i\|_T = (a_{ii}^T(0))^{1/2} \quad i = 1, 2, \dots, K$$

以下では (G.1) ~ (G.4) を仮定する.

(i)  $f(\lambda)$ : continuous とする.

$$\begin{aligned} \lim_{T \rightarrow \infty} D_T E\{(\hat{\beta}_{T, LSE} - \beta)(\hat{\beta}_{T, LSE} - \beta)'\} D_T \\ = 2\pi R(0)^{-1} \int_{-\pi}^{\pi} f(\lambda) dM(\lambda) R(0)^{-1} \end{aligned}$$

(ii)  $f(\lambda)$ : positive continuous とする

$$\begin{aligned} \lim_{T \rightarrow \infty} D_T E\{(\hat{\beta}_{T, BLU} - \beta)(\hat{\beta}_{T, BLU} - \beta)'\} D_T \\ = [(2\pi)^{-1} \int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda)]^{-1} \end{aligned}$$

(iii) LSE が漸近有効になるための必要十分条件は,

□  $M(\lambda)$  が  $0 \leq \lambda \leq \pi$ , 高々  $k$  個の点で増加し,

各点での行列, 階数の総和が  $k$  になる □ こと

である. 但し,  $f(\lambda)$  は positive continuous とする.

次にいくつかの記号・仮定を導入する.

記号

$$m_{p\bar{q}}^T(\lambda) = \frac{(\sum_{t=1}^T x_{tp} e^{-i\lambda t})(\sum_{t=1}^T x_{t\bar{q}} e^{i\lambda t})}{2\pi \|X_p\|_T \|X_{\bar{q}}\|_T}$$

$$M_{p\bar{q}}^T(\lambda) = \int_{-\pi}^{\lambda} m_{p\bar{q}}^T(\omega) d\omega$$

\*

$$M^T(\lambda) = [M_{pq}^T(\lambda)]_{k \times k}$$

$M^T(\lambda)$  を complex matrix measure とみよせば、(G.3) は

$$M^T(\lambda) \xrightarrow{w} M(\lambda) \quad (T \rightarrow \infty)$$

を意味する。(ここで  $\xrightarrow{w}$  は弱収束)。この事実が以下の結果を導く上で、key point になる。

仮定. 必要なるは説明変数の番号を入れ換えることにより、

$$M_{ii}(0+) - M_{ii}(0) > 0, \quad i=1, \dots, m, 0 \leq m \leq k$$

$$M_{ii}(0+) - M_{ii}(0) = 0, \quad i=m+1, \dots, k.$$

とする。

### Theorem 1 (LSE の漸近分散)

以下、(G.1) ~ (G.4) を仮定する。また  $\{\varepsilon_t\}$  は長期記憶モデルにしたがうとする。

(i)  $m = 0$  とする。

Theorem (Grenander) の (i) の等式成立

$$\Leftrightarrow \forall \delta (> 0) \exists c \text{ at}$$

$$(*) \quad \int_{|\lambda| \leq c} f(\lambda) dM_{ii}^T(\lambda) < \delta, \quad i=m+1, \dots, k, \quad \forall T.$$

(ii)  $m > 0$  とする。(\*) を仮定

(a)  $\int_{-\pi}^{\pi} f(\lambda) dM_{ij}^T(\lambda) / T^{2\alpha}$  の極限存在 ( $T \rightarrow \infty$ ),  $1 \leq i, j \leq m$

$\Leftrightarrow \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2\alpha} dM_{ij}^T(\lambda) / T^{2\alpha}$  の極限存在 ( $T \rightarrow \infty$ ),  $1 \leq i, j \leq m$

(b)  $\widehat{D}_T \approx \text{diag}(\|X_1\|_T \cdot T^\alpha, \dots, \|X_m\|_T \cdot T^\alpha, \|X_{m+1}\|_T, \dots, \|X_K\|_T)$

$h_{ij}^{(1)} \approx \lim_{T \rightarrow \infty} \int_{-\pi}^{\pi} |1 - e^{i\lambda}|^{-2\alpha} dM_{ij}^T(\lambda) / T^{2\alpha}, 1 \leq i, j \leq m$

このとき

$$\lim_{T \rightarrow \infty} \widehat{D}_T^{-1} (\widehat{X}_T' \widehat{X}_T) E\{(\widehat{\beta}_{T, LSE} - \beta)(\widehat{\beta}_{T, LSE} - \beta)'\} (\widehat{X}_T' \widehat{X}_T) \widehat{D}_T^{-1} = 2\pi B$$

ここで

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

$$B_1 = [h_{ij}^{(1)}]_{m \times m} \quad B_2 = \left[ \int_{-\pi}^{\pi} f(\lambda) dM_{i+m, j+m}(\lambda) \right]_{(K-m) \times (K-m)}$$

次に BLUE について考える。(G. 1) ~ (G. 4) に加えて以下の条件を課す。

$$(G. 5) \quad \max_{1 \leq k \leq T} X_{ki}^2 / a_{ii}^T(0) = o(1/T^\delta), \delta > 1 - 2\alpha, i = 1, \dots, K.$$

### Theorem 2. (BLUE の漸近分散. I)

次の仮定をおく。

$$0 < M_{ii}(0+) - M_{ii}(0) < 1, \quad i = 1, \dots, m$$

$$\int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda) \text{ が nonsingular}$$

このとき (G. 1) ~ (G. 5) の下で、Theorem (Grenander) の (ii)

の等式が成立する。

ところで

$$M_{\lambda\lambda}(0+) - M_{\lambda\lambda}(0) = 1 \quad \text{for } \lambda \quad (1 \leq \lambda \leq m)$$

のときは、 $\int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda)$  は singular になる。一般的な場合は未解決である。ここでは典型的な例として多項式を考える。すなわち

$$X_{\lambda\lambda} = \lambda^{p-1}, \quad \lambda = 1, \dots, p, \quad 0 \leq p \leq m$$

とする。したがって

$$M_{\lambda\lambda}(0+) - M_{\lambda\lambda}(0) = 1, \quad \lambda = 1, \dots, p.$$

$$M_{\lambda\lambda}(0+) - M_{\lambda\lambda}(0) < 1, \quad \lambda = p+1, \dots, m.$$

とする。

### Theorem 3 (BLUEの漸近分散 II)

$$\tilde{D}_T \propto \text{diag}(\|X_1\|_T/T^\alpha, \dots, \|X_p\|_T/T^\alpha, \|X_{p+1}\|_T, \dots, \|X_m\|_T)$$

上述の仮定および (G.1) ~ (G.5) の下で、

$$\begin{aligned} \lim_{T \rightarrow \infty} \tilde{D}_T E\{(\hat{\beta}_{T, BLUE} - \beta)(\hat{\beta}_{T, BLUE} - \beta)'\} \tilde{D}_T \\ = 2\pi W \end{aligned}$$

ここで

$$W = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$



$$W_1 = \{W_{\lambda, d}^{(1)}\}_{p \times p} \quad W_{\lambda, d}^{(1)} = \frac{\Gamma(\lambda-d)\Gamma(d-d)\{(2\lambda-1)(2d-1)\}^{1/2}}{f^*(0)\Gamma(\lambda-2d)\Gamma(d-2d)(\lambda+d-1-2d)}$$

$$W_2 = \left\{ \int_{-\pi}^{\pi} f(\lambda) dM_{\lambda+p, d+p}(\lambda) \right\}_{(k-p) \times (k-p)}$$

以上の結果より  $\angle SE$  の漸近有効性について、Theorem (Grenander) の (iii) を一般化して、次の定理を得る。

### Theorem X (LSE の漸近有効性)

次の仮定をおく。

$$M_{\lambda, i}(0+) - M_{\lambda, i}(0) = 0, \quad 1 \leq i \leq k$$

$$\forall \delta > 0 \exists c \text{ st}$$

$$\int_{|\lambda| \leq c} f(\lambda) dM_{\lambda, i}^T(\lambda) < \delta, \quad 1 \leq i \leq k, \quad \forall T.$$

このとき (G.1) ~ (G.5) の下で、 $\angle SE$  が漸近有効になるための必要十分条件は、『 $M(\lambda)$  が  $0 < \lambda \leq \pi$  の高々  $k$  個の点で増加し、各点での行列の階数の総和が  $k$  となる』ことである。

Theorem X の直観的解釈は次の様になる。 $M(\lambda)$  が  $\lambda=0$  で増加するとき、 $f(\lambda)$  も  $\lambda=0$  において発散するので、

$\angle SE$  は説明変数と誤差項のスペクトルを判別できず、漸近有効にはなるない。しかし  $M(\lambda)$  が  $0 < \lambda \leq \pi$  で増加する場合

合には、 $f(x)$ は半閉区間  $(0, \pi]$  で有限値を取るので、 $\angle SE$ は  $M(x)$ と  $f(x)$ を分離することができ、漸近有効となる。

ここで Theorem 1~4の含意を明確にするため、ひとつの例をあげる。

### Example

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_t$$

ここで

$$x_{t1} = \tau_{10} + \tau_{11} \cos V_1 t + \tau_{12} \cos V_2 t$$

$$x_{t2} = \tau_{20} + \tau_{21} \cos V_1 t + \tau_{22} \cos V_2 t$$

$$V_i \neq 0, i=1, 2, \quad V_1 \neq V_2.$$

このとき

$$R(h) = M_0 + \cos V_1 M_1 + \cos V_2 M_2$$

但し、 $M_0 = \Gamma \tilde{\tau}_0 \tilde{\tau}_0' \Gamma$ ,  $M_i = \Gamma \tilde{\tau}_i \tilde{\tau}_i' \Gamma$ ,  $i=1, 2$ .

$$\Gamma = \text{diag} \left( \frac{1}{\sqrt{\tau_{10}^2 + (\tau_{11}^2 + \tau_{12}^2)/2}}, \frac{1}{\sqrt{\tau_{20}^2 + (\tau_{21}^2 + \tau_{22}^2)/2}} \right)$$

$$\tilde{\tau}_i = (\tau_{i1}, \tau_{i2})', \quad i=0, 1, 2.$$

いま  $\angle SE$ の漸近相対効率を

$$e(d) = \lim_{T \rightarrow \infty} \frac{\det [E(\hat{\beta}_{T, BLU} - \beta)(\hat{\beta}_{T, BLU} - \beta)']}{\det [E(\hat{\beta}_{T, LSE} - \beta)(\hat{\beta}_{T, LSE} - \beta)']}$$

とする。  $\tau_i$  について以下の三例を考える。各々の例に

おいて  $e(d)$  は互いに異なる値をとる.

$$(i) \quad \tilde{\tau}_0 = (0, 0)', \quad \tilde{\tau}_1 = (1, 0)', \quad \tilde{\tau}_2 = (0, 1)'$$

Th. 1. (i), Th. 2, Th. 4 より

$$e(d) = 1, \quad \forall d.$$

$$(ii) \quad \tilde{\tau}_0 = (1, 0)', \quad \tilde{\tau}_1 = (1, 0)', \quad \tilde{\tau}_2 = (0, 1)'$$

Th. 1. (ii), Th. 2 より

$$e(d) = 0, \quad \forall d.$$

$$(iii) \quad \tilde{\tau}_0 = (1, 0)', \quad \tilde{\tau}_1 = (0, 1)', \quad \tilde{\tau}_2 = (0, 0)'$$

Th. 1. (ii), Th. 3 より

$$0 < e(d) = (1+2d)P(1+d)P(2-2d)/P(1-d) < 1, \quad \forall d.$$

なお  $f(\lambda)$  が positive continuous の場合は、漸近相対効率  $e$  を  $e(d)$

と同様に定義すれば、Theorem (Grenander) より各例について

$$(i) \quad e = 1 \quad (ii) \quad 0 < e < 1 \quad (iii) \quad e = 1$$

を得る。

### 2 SE の漸近分布

#### Theorem 5

次の仮定をおく

$\{\varepsilon_t\}$  強定常過程  $E|\varepsilon_t|^k < \infty, \forall k$

$\{\varepsilon_t\}$  の white noise processes の任意の次数のキュムラント

は絶対総和可能

さらに Th. 1 (ii) と同様の仮定、および以下の二条件を導入する。

$$\max_{|\lambda| \leq \delta} (m_{ii}^T(\lambda))^{1/2} = o(1/T^{1/2 + d/2}), \quad i=m+1, \dots, k, \quad \forall \delta > 0$$

$$\int_{-\pi}^{\pi} (m_{ii}^T(\lambda))^{1/2} d\lambda = o(1/T^{1/2 + d/2}), \quad i=1, \dots, k.$$

このとき

$$\hat{D}_T^{-1}(\tilde{X}_T' \tilde{X}_T)(\hat{\beta}_T - \beta) \xrightarrow{L} N(0, 2\pi B) \quad (T \rightarrow \infty)$$

ここで  $\xrightarrow{L}$  は法則収束とする。

### Remark

具体的な例として

$$x_{ti} = t^{i-1}, \quad 1 \leq i \leq m,$$

$$x_{ti} = \cos \nu_i t, \sin \nu_i t, \text{ あるいは周期関数 } m+1 \leq i \leq k$$

ただし  $\nu_i \neq 0$  とすれば

$$\max_{|\lambda| \leq \delta} (m_{ii}^T(\lambda))^{1/2} = O(1/T^{1/2}), \quad i=m+1, \dots, k, \quad \forall \delta > 0$$

$$\int_{-\pi}^{\pi} (m_{ii}^T(\lambda))^{1/2} d\lambda = O(\log T/T), \quad i=1, \dots, k.$$

となり、Theorem 5 の仮定を満足する。

### 3. 今後の問題

回帰モデルを含め、今後解決すべき問題を二、三列举する。

(1) 回帰モデルにおいて  $\angle SE$  が漸近有効でないとき、

$\angle SE$ より相対効率の良い推定量は存在するか？（ただしBLUEは除外して）

例  $y_t = \beta_1 + \varepsilon_t$

(i)  $\{\varepsilon_t\}$ のスペクトル密度が  $f(\lambda) = 1 / |1 - e^{i\lambda}|^{2d}$

つまり  $f^*(\lambda) \equiv 1$  のときのBLUEは、真のスペクトル密度が  $f(\lambda) = f^*(\lambda) / |1 - e^{i\lambda}|^{2d}$ ,  $f^*(\lambda)$ は任意の正の連続関数であっても、漸近的には有効となる。

(Adenstedt (1974))。しかし実際には  $d$  が未知なので何らかの推定量で代替しなくてはならない。

(ii) 任意の  $M$ -estimator の漸近相対効率は  $\angle SE$  のそれに等しい。(Beran & Künsch (1985)) 従って  $M$ -estimator によって  $\angle SE$  を改善することはできない。

(iii)  $h(t) : [0, 1] \rightarrow \mathbb{R}$  有界変動な実関数

$h(t)$  を taper として用いると、

$$\lim_{T \rightarrow \infty} \frac{T^{1-2d} \text{Var} \left( \sum_{t=1}^T h(t/T) y_t / \sum_{t=1}^T h(t/T) \right)}{T^{1-2d} \text{Var} \left( \sum_{t=1}^T y_t / T \right)}$$

$$= \frac{\int_0^1 \int_0^1 h(x) h(y) |x-y|^{2d-1} dx dy}{\left\{ \int_0^1 h(x) dx \right\}^2 \int_0^1 \int_0^1 |x-y|^{2d-1} dx dy}$$

上式は 算術平均（この場合の  $\angle SE$ ）の分散と、taperを用いた線形不偏推定量の分散の比の極限である。これを1より小さくする  $h(t)$  は存在するか？

(2) 検定に関しては local alternative などの問題がある  
たとえば

$$H_0: d = d_0 \quad H_1: d = d_0 + c/\sqrt{T}$$

とした場合、尤度比の性質、contiguity, LAN族  
に属するかが問題となる。特に  $d_0 = 0$  のときは、  
短期記憶モデル VS 長期記憶モデルの検定となる。

(3) 最後に、真のスペクトル密度が想定するパラメトリック  
モデルに属さない incorrect model の問題がある。  
例.

$$\text{真のスペクトル密度 } f(\lambda) = f^*(\lambda) / |1 - e^{i\lambda}|^{2\alpha}, 0 < \alpha < 1/2$$

$$\text{想定するモデル } f(\lambda) = |\theta(e^{i\lambda})|^2 / |\phi(e^{i\lambda})|^2$$

ARMAモデル

このとき MLE 等はどのような漸近的性質を持つか?

「以上」

#### References

- Adenstedt, R.K. (1974). On large-sample estimation for the mean of a stationary random sequence. Ann. Statist. 2 1095-1107.
- Beran, J. (1989). A test of location for data with slowly decaying serial correlations. Biometrika 76 261-269.
- Beran, J. and Kunsch, H. (1985). Location estimators for processes with long range dependence. Res. Rep. 40 Seminar fur Statistik. ETH. Zurich.

- Grenander, U. (1954). On estimation of regression coefficients in the case of an autocorrelated disturbance. Ann. Math. Statist. 25 252-272.
- Samarov, A. & Taqqu, M. S. (1988). On the efficiency of the sample mean in long-memory noise. J. Time Series Analysis 9 191-200.
- Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. Ann. Statist. 16 791-807.